Development of an Oral Benchmark Dose for Sulfolane Based on Data from a 90-Day Drinking Water Study Conducted by Huntingdon Life Sciences (2001)

AUGUST 18, 2010

ToxStrategies

Innovative solutions
Sound science

Development of an Oral Benchmark Dose for Sulfolane Based on Data from a 90-Day Drinking Water Study Conducted by Huntingdon Life Sciences (2001)

AUGUST 18, 2010

PREPARED FOR:
Flint Hills Resources
North Pole, Alaska Refinery

PREPARED BY:
ToxStrategies, Inc.
3420 Executive Center Drive
Suite 114
Austin, Texas
USA

HLS Study Background

Huntingdon Life Sciences (HLS, 2001) conducted a GLP certified subchronic (90-day) drinking water study for sulfolane in CD (Sprague Dawley) rats. In this study, male and female rats were exposed to concentrations of 0, 25, 100, 400, or 1,600 mg/L sulfolane (2.1, 8.8, 35, and 132 mg/kg/day in male rats and 2.9, 10.6, 42 and 191 mg/kg/day in female rats) *ad libitum*. Ten male and 10 female rats were exposed to each drinking water concentration. For quality assurance, samples of each sulfolane formulation prepared for administration were analyzed for achieved concentration at weeks 1, 6 and 12 of the study. The animals were thoroughly examined for signs of adverse health effects. Examinations included: food and water consumption, bodyweight, organ weights, functional observations (e.g. reflexes, grooming, motor activity), hematological evaluations, blood chemistry, gross pathology, and histopathological examination of 13 of the major organs (adrenals, brain, femur, heart, ileum, kidneys, liver, lungs, mammary area, spinal cord, stomach, thyroid, and uterus).

The exposure was described as well tolerated and the study authors noted two treatment-related effects following oral exposure to sulfolane. Specifically, male rats exhibited treatment related effect in kidneys involving both cortical cell basophilia and hyaline droplets. This form of renal toxicity in male rats is well documented and involves chemical-induced inhibition of alpha-2u-globulin catabolism leading to accumulation in secondary lysosomes that appear transparent (hyaline-like) via microscopy. This protein appears to be specific to male rats and thus the hyaline droplets are not considered to be relevant to humans (U.S. EPA, 1991; Hard et al., 1993; HLS, 2001). In fact, USEPA has concluded that alpha-2u-globulin hyaline droplet formation is unique to male rats and is probably not relevant to humans for purposes of risk assessment (U.S. EPA, 1991). The basophilia observed likely relates to cell proliferation (tubular regeneration) subsequent to alpha-2u-globulin accumulation. Indeed the HLS study authors attributed this and the presence of granular casts (cell debris) in the highest male dose group to "hydrocarbon nephropathy" due to alpha-2u-globulin accumulation. According to Hard et al. (1993), "granular casts stain positive for alpha-2u-globulin, indicating probable derivation of debris from cells that had accumulated this protein." Based on the aforementioned considerations, these

endpoints were not considered relevant for the assessment of human health risk posed by sulfolane.

The other effect considered to be treatment-related by the HLS study authors was a decrease in lymphocyte, monocyte and large unstained cell counts (and a concomitant decrease in total white blood cell (WBC)/leukocyte counts) in female rats administered 100, 400, or 1600 mg/l (10.6, 42 and 191 mg/kg/day, respectively), though the study authors concluded that these effects did not follow a strong trend with dose. Additionally, the study authors noted that there was no evidence of any chronic inflammatory change or compromised immune function in females, nor were there any effects on bone marrow, thymus or spleen that might account for reduced numbers of white blood cells. As such, the study authors concluded that the toxicological significance of the effects on white blood cells was unclear. Further complicating the understanding of the toxicological significance of these findings is the fact that these effects were not observed in male rats. Nonetheless, in the interest of being conservative (i.e., health protective), this endpoint was considered in the development of an oral toxicity benchmark for sulfolane as described in the sections which follow.

Statistical Analyses

The HLS study investigators reported statistically significant decreases in white blood cell, lymphocyte, monocyte, and large unstained cell counts in female rats given 100 mg/l (10.6 mg/kg/day) or more based on Williams' Test. However, as already noted above, even though these decreases were statistically significant relative to the concurrent control animals, the HLS study investigators concluded that the toxicological significance of the effects on white blood cells was unclear due to the lack of evidence of any chronic inflammatory change or compromised immune function in female rats, as well as lack of any effects on bone marrow, thymus or spleen that might account for reduced numbers of white blood cells.

To more closely examine these hematological effects, we performed a series of statistical analyses for total white blood cell (WBC) counts, as well as on counts of the various WBC

components – including lymphocytes, basophils, monocytes, and large unstained cells (LUCs). As is evident in **Table 1**, pair wise t-tests on the means of each dose group relative to controls indicated a statistical significant decrease (p-value ≤ 0.05) at the three highest dose levels tested (100, 400, and 1600 mg/L or 10.6, 42 and 191 mg/kg/day). In addition, trend tests (1- and 2-sided Joncheere's test; Jonckheere, 1954) demonstrated a statistically significant decreasing trend for total WBC counts, as well as counts of all four WBC components, as dose increased.

Table 1. Statistical Summary of HLS Leukocyte Data

Endpoint	Cell Count x 10^9, Mean (s.d.) t-test					Trend Tests	
Dose, mg/kg/day (n)	0 (10)	2.9 (10)	10.6 (9)	42.0 (9)	191.1 (10)	Jonckheere[1]	Jonckheere[2]
Total WBCs[3]	7.97 (2.213)	7.763 (2.653)	5.41 (1.392)	5.53 (1.756)	4.54 (1.019)		
p-values		0.76	0.008	0.016	0.001	0.00013	0.00002
Lymphocytes	6.98 (2.146)	6.36 (2.452)	4.39 (1.308)	4.63 (1.564)	3.73 (0.941		
p-values		0.56	0.006	0.014	0.001	0.00006	0.00001
Basophiles	0.01 (0.006)	0.01 (0.006)	0 (0.005)	0 (0.007)	0 (0.004)		
p-values		0.44	0.018	0.046	0.001	0.00062	0.00011
Monocytes	0.22 (0.08)	0.23 (0.119)	0.13 (0.053)	0.13 (0.040)	0.10 (0.040)		
p-values		0.94	0.012	0.008	0.001	0.00018	0.00003
LUCs	0.11 (0.040)	0.11 (0.056)	0.06 (0.023)	0.06 (0.026)	0.04 (0.019)		
p-values		0.93	0.011	0.007	0.001	0.00002	0.000003

[1] Jonckheere 2-sided test (http://www.biostat.wustl.edu/archives/html/s-news/2000-10/msg00126.html)
[2] Jonckheere 1-sided test for decreasing trend (http://tolstoy.newcastle.edu.au/R/help/06/06/30112.html)
[3] WBC-white blood cells; LUCs-large unstained cells

In addition to statistical significance, it is also important to determine if the effects are biologically meaningful. To accomplish this, the white blood cell counts in the HLS sulfolane drinking water study were compared to historical control data from the same HLS laboratory. Blood cell counts from individual female rats in the HLS sulfolane drinking water study were converted to incidence counts using the historical control ranges. These incidence counts were then compared using Fisher's Exact Test with Holm's correction for multiple comparisons. There were no significant differences between dose groups and the control. This finding suggests that though statistical analyses indicate a treatment-related

decrease in WBC counts at the three highest dose levels tested in female rats, the effects appear to be subtle as they are not outside of the historical control range for the specific species, strain, gender, age, and laboratory. This conclusion is consistent with that of the HLS study authors conclusion that the toxicological significance of the WBC effects was unclear due to the lack of evidence of any chronic inflammatory change or compromised immune function in female rats, as well was lack of any effects on bone marrow, thymus or spleen that might account for reduced numbers of white blood cells.

Development of Toxicity Benchmarks Based on the HLS Study (2001)

Risk assessors have argued for at least two decades that the LOAEL/NOAEL[1] approach is an inferior risk assessment approach because it is limited to the doses tested, does not appropriately address study size, does not allow for direct comparisons across studies and endpoints based on a common response level (e.g. 10% increased risk), and can inappropriately reward poorer studies with less statistical power to detect effects resulting in higher LOAEL and NOAEL values (Crump, 1984; Leisenring and Ryan, 1992; Gaylor et al., 1998; Allen et al., 1998; U.S. EPA, 2000, 2002b). In contrast, benchmark dose (BMD) modeling has been recognized as the preferred alternative because it takes into account the shape of the dose-response curve, the confidence limits reflect the size of the study, and allows comparison of comparable results across studies and endpoints at any response level (e.g. 10% increased risk) (Allen et al., 1998; U.S. EPA, 2000, 2002b). As such, the preferred approach for identifying a health protective point of departure (POD) from the HLS (2001) study is to use BMD modeling approaches. Such approaches provide a mathematically and biologically supportable approach to establishing PODs. The BMD modeling of the HLS data are described below.

Based on apparent trend in the data and aforementioned statistical tests, we modeled the total WBC, lymphocyte, monocytes, and large unstained cells (LUC) cell count data (a continuous variable) using the U.S. EPA Benchmark Dose Software (BMDS). Basophils were

---

[1] LOAEL = lowest observable adverse effect level; NOAEL = no observable adverse effect level

not modeled because the mean values for the control and four dose groups were 0.01, 0.01, 0.00, 0.00 and 0.00×10^9 cells/L.

Initially, none of the models in the BMDS were able to reasonably fit the total WBC, lymphocyte, monocyte, or LUC cell count data. In such instances, risk assessors sometimes drop the highest dose in the study and remodel the data to improve the model fit. However, this should only be done if there is evidence of lethality at the highest dose level or there is evidence of a plateau in the toxic response at the highest dose level. As this is not the case with the HLS dataset, there is no scientifically supportable basis for dropping the highest dose level solely for model fitting purposes. However, the dose spacing in the HLS study was such that the two lower doses covered only a small proportion (10.6 / 191.1 = 5.5%) of the total dose range, and thus the higher doses unduly influence the model fit. A scientifically supportable approach for addressing a situation like this is to log transform the doses. By log transforming the doses, the lower doses take on a more even spacing (2.5/5.3 = 47%), and lowers the influence of the high dose without arbitrarily dropping data points.

Log transformation of dose has previously been used to model a reduction in lymphocytes in the toxicological review of the noncancer effects of benzene (U.S. EPA, 2002a). Therein, reference concentration (RfC) and reference dose (RfD) values were established using lymphocyte count data from humans exposed to benzene via inhalation. All of the continuous models from the BMDS produced poor fits to the benzene data, in part, due to the supralinear response pattern. Therefore, the EPA log transformed the doses, ln(dose+1), and remodeled the untransformed responses against transformed doses. This resulted in models that fit the data, and EPA stated, "the linear model was selected because it is the most parsimonious." The resulting BMD and BMDL values were then converted back to arithmetic dose as follows: $e^{\ln(dose+1)} - 1$.

Applying the same approach to the total WBC and lymphocyte data from the HLS study, along with U.S. EPA's default approach of using a BMR based on 1 standard deviation from

the concurrent study controls and fitting a linear model, we computed default $BMDL_{1SD}$ values of 11.9 and 14.5 mg/kg/day based on total WBC and lymphocyte counts, respectively (with acceptable p-values)[2]. Using the same approach, models were unable to achieve reasonable fits to the cell counts for monocytes or LUC. As such, these two endpoints were not considered further in our analyses. This was determined to be inconsequential since all hematological endpoints exhibited the same LOAEL in female rats and total WBC counts were observed to be most dependent on lymphocyte counts in the HLS dataset for female rats.

When modeling continuous datasets, it has also been suggested that the data can be modeled using the historical standard deviation from control (i.e. untreated) animals along with the concurrent control mean data from the study of interest (U.S. EPA, 2000). In contrast to the limited number of female rats in the concurrent control group in the HLS sulfolane study (N=10), historical control data can provide a better indication of the true variability of a given biochemical or toxicological endpoint. As such, we obtained historical control hematology data for 393 female CD Sprague-Dawley rats of 16-21 weeks of age from HLS. This historical control data is ideal because it comes from the same species, strain, sex, and age group of animals from the same laboratory and time period as in the sulfolane study; and moreover, because these data come from the same HLS laboratory, the total WBC and lymphocyte counts were most likely obtained using the same collection and analytical techniques as were used in the sulfolane study. Therefore, the HLS historical control data provides a much more robust dataset for establishing the normal range of variability for the endpoints of interest. Given this, the WBC and lymphocyte standard deviations from the concurrent HLS control animals (2.213 and 2.146, respectively) were replaced with those from the historical HLS dataset (2.626 and 2.290, respectively), and BMD modeling was again performed with the BMR set to 1 standard deviation. In accordance with this approach, the $BMDL_{1SDh}$ values for total WBC and lymphocytes were determined using a linear model. Because the standard deviations from the historical data were slightly greater than those of the HLS study, the resulting BMD and BMDL values

---

[2] We also considered that the data may be lognormally distributed; however, Pearsons's index of skewness did not indicate that the data were strongly skewed.

increased slightly compared to those calculated using the concurrent control standard deviation. These BMDL values were **15.1** and **16.0** mg/kg/day for WBC and lymphocytes, respectively (**Table 2**). However, because this historical standard deviation is drawn from a much larger sample size (393 vs 10), the historical standard deviation is a more representative measure of the true variability in total WBC and lymphocyte counts. As such, the BMD and BMDL values derived based on the historical control standard deviation are more defensible scientifically than are those derived using the standard deviation based on the limited number of concurrent study controls. The BMD model output for this analysis is provided in Attachment A.

Table 2. Summary of Benchmark Modeling Results Based on a Linear Model

Model Parameter	Reduced Leukocyte Count	
	WBC	**Lymphocytes**
p-values	0.1677	0.158
scaled residual	0.168	0.232
BMD^{1}, ln(dose + 1)	4.22	4.34
BMDL, ln(dose + 1)	2.78	2.83
BMD	67.03	75.71
BMDL	15.12	15.95

[1] Because the doses were log transformed, the BMD and BMDL values reported in the BMD software output were ln(dose + 1) and were manually converted back to arithmetic scale for reporting BMD and BMDL

Alternative BMD Models and Modeling Approaches

The BMD and BMDL values reported in **Table 2** reflect results obtained when fitting a linear model and were ultimately chosen as the PODs based on considerations of model fit, as well as on the basis of parsimony. This is consistent with the approach taken by U.S. EPA in their selection of results from the linear model in their BMD modeling of the effects of benzene on lymphocytes (U.S. EPA, 2002a). A comparison between these BMD and BMDL values to those obtained with other models and modeling approaches is provided in **Attachment B.**

Recommend POD for Risk Assessment

The BMD analyses summarized above (and in **Attachments A and B**) provide a narrow range of potential POD values that can be used to develop an oral RfD for sulfolane, suggesting generally good agreement between the different modeling approaches. Changes in WBC counts appear to be a sensitive indicator of sulfolane exposure; however as already discussed above, the adversity of this effect remains to be demonstrated. Nonetheless, in the interest of being conservative (i.e., health protective), these endpoints were treated as if they were "adverse" for purposes of developing a POD for risk assessment purposes even though this has not been clearly demonstrated by the available data. Given the much larger sample size for historical vs concurrent control animals (393 vs 10, respectively), the historical standard deviation is believed to be a better measure of variability in leukocyte counts in untreated animals. As already noted, the linear model provides the best fit and is the most parsimonious and, as such, BMDL values from this model were selected as the PODs. Examination of **Table 2** indicates that the BMDL from the linear model for total WBC was slightly lower than that for lymphocytes. Therefore, the recommended POD is based on the $BMDL_{1SDh}$ for decreases in total WBC of **15.1 mg/kg/day**.

References

Andersen et al. (1976) Res. Communication in Chemical Path and Pharm. 15, 571-

Andersen et al. (1977) Toxicol. Applied Pharm. 40, 472-

Allen et al. (1998) Regulatory Tox and Pharm. 28, 38-

Crump (1984) Fundamental Appl. Toxicol. 4: 854-

Gaylor et al. (1998) Regulatory Toxicol. Pharmacol. 28: 150-

Hard et al., (1993). Hazard Evaluation of Chemicals That Cause Accumulation of a2u-globulin, Hyaline Droplet Nephropathy, and Tubule Neoplasia in the Kidneys of Male Rats. Environmental Health Perspectives. 99, 313-349

Huntingdon Life Sciences (HLS). 2001. Sulfolane Toxicity Study by Oral Administration via the Drinking Water Pathway to CD Rats for 13 Weeks. Huntingdon Life Sciences Ltd. Huntingdon, England.

Jonckheere, 1954

Leisenring and Ryan (1992) Regulatory Toxicol. Pharmacol. 15: 161-

Suckow et al. (2006). The Laboratory Rat. Elsevier Academic Press, Burlington, MA.

U.S. EPA (1991). Report of the EPA Peer Review Workshop on Alpha-2u-globulin: Association with Chemically Induced Renal Toxicity and Neoplasia in the Male Rat. Risk Assessment Forum. EPA/625/3-91/019F. http://epa.gov/osa/raf/publications/alpha2u-globulin.htm

U.S. EPA. (2000) Benchmark Dose Technical Guidance Document, Risk Assessment Forum, Washington, DC, EPA/630/R-00/001.

U.S. EPA (2002a). IRIS Toxicological Review of Benzene (Noncancer Effects). U.S. Environmental Protection Agency, Office of Research and Development, National Center for Environmental Assessment, Washington Office, Washington, DC, EPA/635/R-02/001F, 2002.

U.S. EPA (2002b). A Review of the Reference Dose and Reference Concentration Processes. U.S. Environmental Protection Agency, Risk Assessment Forum, Washington, DC, EPA/630/P-02/002F, 2002.

Attachment A

BMD Model Output – Linear Model Results

```
================================================================
        Polynomial Model. (Version: 2.13;   Date: 04/08/2008)
        Input Data File: C:\USEPA\BMDS21\Data\lin_HLS-WBC_logdose-historicalHLS-
SD_Setting.(d)
        Gnuplot Plotting File:  C:\USEPA\BMDS21\Data\lin_HLS-WBC_logdose-
historicalHLS-SD_Setting.plt
                                        Wed Jun 02 14:30:09 2010
================================================================
```

BMDS Model Run
~~~~~~

   The form of the response function is:

   Y[dose] = beta_0 + beta_1*dose + beta_2*dose^2 + ...

   Dependent variable = Response
   Independent variable = Dose
   Signs of the polynomial coefficients are not restricted
   The variance is to be modeled as Var(i) = exp(lalpha + log(mean(i)) * rho)

   Total number of dose groups = 5
   Total number of records with missing values = 0
   Maximum number of iterations = 250
   Relative Function Convergence has been set to: 1e-008
   Parameter Convergence has been set to: 1e-008

                    Default Initial Parameter Values
                          lalpha =   1.41295
                            rho =   0
                         beta_0 =   7.97188
                         beta_1 =   -0.684223

           Asymptotic Correlation Matrix of Parameter Estimates

                  lalpha   rho   beta_0   beta_1

      lalpha   1   -0.99   0.11   -0.13

         rho   -0.99   1   -0.11   0.13

      beta_0   0.11   -0.11   1   -0.91

      beta_1   -0.13   0.13   -0.91   1

                          Parameter Estimates

                                                   95.0% Wald Confidence
Interval
       Variable   Estimate   Std. Err.   Lower Conf. Limit   Upper Conf.
Limit
         lalpha   -4.6222   1.92724   -8.39952   -
0.844874
            rho   3.21044   1.0586   1.13562
```

```
5.28525
        beta_0              7.8725          0.56274              6.76955
8.97545
        beta_1             -0.64517        0.137292            -0.914258            -
0.376082
```

Table of Data and Estimated Values of Interest

Dose	N	Obs Mean	Est Mean	Obs Std Dev	Est Std Dev	Scaled Res.
0	10	7.97	7.87	2.63	2.72	0.113
1.361	10	7.63	6.99	2.65	2.25	0.893
2.451	9	5.41	6.29	1.39	1.9	-1.39
3.761	9	5.53	5.45	1.76	1.51	0.168
5.258	10	4.54	4.48	1.11	1.1	0.172

Model Descriptions for likelihoods calculated

Model A1: $Yij = Mu(i) + e(ij)$
        $Var\{e(ij)\} = Sigma^2$

Model A2: $Yij = Mu(i) + e(ij)$
        $Var\{e(ij)\} = Sigma(i)^2$

Model A3:   . $Yij = Mu(i) + e(ij)$
        $Var\{e(ij)\} = exp(lalpha + rho*ln(Mu(i)))$
    Model A3 uses any fixed variance parameters that
    were specified by the user

Model R: $Yi = Mu + e(i)$
        $Var\{e(i)\} = Sigma^2$

                Likelihoods of Interest

Model	Log(likelihood)	# Param's	AIC
A1	-55.270870	6	122.541740
A2	-49.849681	10	119.699362
A3	-50.012552	7	114.025103
fitted	-52.540846	4	113.081693
R	-65.052796	2	134.105592

                Explanation of Tests

Test 1: Do responses and/or variances differ among Dose levels?
         (A2 vs. R)
Test 2: Are Variances Homogeneous? (A1 vs A2)
Test 3: Are variances adequately modeled? (A2 vs. A3)
Test 4: Does the Model for the Mean Fit? (A3 vs. fitted)
(Note: When rho=0 the results of Test 3 and Test 2 will be the same.)

                Tests of Interest

Test	-2*log(Likelihood Ratio)	Test df	p-value
Test 1	30.4062	8	0.0001791

Test 2	10.8424	4	0.02839
Test 3	0.325741	3	0.9551
Test 4	5.05659	3	0.1677

The p-value for Test 1 is less than .05. There appears to be a
difference between response and/or variances among the dose levels
It seems appropriate to model the data

The p-value for Test 2 is less than .1. A non-homogeneous variance
model appears to be appropriate

The p-value for Test 3 is greater than .1. The modeled variance appears
 to be appropriate here

The p-value for Test 4 is greater than .1. The model chosen seems
to adequately describe the data

Benchmark Dose Computation

Specified effect =   1

Risk Type   =   Estimated standard deviations from the control mean

Confidence level =   0.95

            BMD =   4.21778

           BMDL =   2.78079

Linear Model with 0.95 Confidence Level

Lymphocytes, Log Dose, Historical HLS SD - Linear

```
====================================================================
          Polynomial Model. (Version: 2.13;  Date: 04/08/2008)
          Input Data File: C:\USEPA\BMDS21\Data\lin_HLS-Lymphocytes-logdose-
historicalHLS-SD_Setting.(d)
          Gnuplot Plotting File:  C:\USEPA\BMDS21\Data\lin_HLS-Lymphocytes-logdose-
historicalHLS-SD_Setting.plt
                                        Wed Jun 02 14:49:52 2010
====================================================================

 BMDS Model Run
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
```

The form of the response function is:

$Y[dose] = beta\_0 + beta\_1*dose + beta\_2*dose^2 + \ldots$

Dependent variable = Response
Independent variable = Dose
Signs of the polynomial coefficients are not restricted
The variance is to be modeled as $Var(i) = exp(lalpha + log(mean(i)) * rho)$

Total number of dose groups = 5
Total number of records with missing values = 0
Maximum number of iterations = 250
Relative Function Convergence has been set to: 1e-008
Parameter Convergence has been set to: 1e-008

```
                    Default Initial Parameter Values
                         lalpha =      1.19837
                            rho =            0
                         beta_0 =      6.83072
                         beta_1 =    -0.628439
```

```
          Asymptotic Correlation Matrix of Parameter Estimates
```

	lalpha	rho	beta_0	beta_1
lalpha	1	-0.99	0.16	-0.19
rho	-0.99	1	-0.16	0.19
beta_0	0.16	-0.16	1	-0.91
beta_1	-0.19	0.19	-0.91	1

```
                         Parameter Estimates
```

			95.0% Wald Confidence Interval	
Variable	Estimate	Std. Err.	Lower Conf. Limit	Upper Conf. Limit
lalpha	-4.2279	1.68468	-7.52982	-0.925975
rho	3.18622	1.02458	1.17808	

```
5.19435
          beta_0          6.68827         0.512056            5.68466
7.69189
          beta_1         -0.574725        0.123487           -0.816754        -
0.332695
```

           Table of Data and Estimated Values of Interest

    Dose   N   Obs Mean   Est Mean   Obs Std Dev   Est Std Dev   Scaled Res.
    ------   ---   --------   --------   -----------   -----------   ----------

        0   10   6.98   6.69   2.29   2.49   0.37
    1.361   10   6.36   5.91   2.45   2.04   0.702
    2.451   9   4.39   5.28   1.31   1.71   -1.56
    3.761   9   4.63   4.53   1.56   1.34   0.232
    5.258   10   3.73   3.67   0.941   0.957   0.21

   Model Descriptions for likelihoods calculated

   Model A1:   Yij = Mu(i) + e(ij)
              Var{e(ij)} = Sigma^2

   Model A2:   Yij = Mu(i) + e(ij)
              Var{e(ij)} = Sigma(i)^2

   Model A3:   Yij = Mu(i) + e(ij)
              Var{e(ij)} = exp(lalpha + rho*ln(Mu(i)))
       Model A3 uses any fixed variance parameters that
       were specified by the user

   Model R: Yi = Mu + e(i)
              Var{e(i)} = Sigma^2

                     Likelihoods of Interest

            Model   Log(likelihood)   # Param's   AIC
             A1   -50.120881   6   112.241762
             A2   -44.447695   10   108.895390
             A3   -44.704461   7   103.408922
           fitted   -47.302522   4   102.605045
             R   -60.319315   2   124.638631

                       Explanation of Tests

   Test 1:   Do responses and/or variances differ among Dose levels?
             (A2 vs. R)
   Test 2:   Are Variances Homogeneous? (A1 vs A2)
   Test 3:   Are variances adequately modeled? (A2 vs. A3)
   Test 4:   Does the Model for the Mean Fit? (A3 vs. fitted)
   (Note: When rho=0 the results of Test 3 and Test 2 will be the same.)

                       Tests of Interest

     Test   -2*log(Likelihood Ratio)   Test df   p-value

     Test 1   31.7432   8   0.0001035
```

Test 2	11.3464	4	0.02294
Test 3	0.513532	3	0.9159
Test 4	5.19612	3	0.158

The p-value for Test 1 is less than .05. There appears to be a
difference between response and/or variances among the dose levels
It seems appropriate to model the data

The p-value for Test 2 is less than .1. A non-homogeneous variance
model appears to be appropriate

The p-value for Test 3 is greater than .1. The modeled variance appears
 to be appropriate here

The p-value for Test 4 is greater than .1. The model chosen seems
to adequately describe the data

               Benchmark Dose Computation

Specified effect =   1

Risk Type   =   Estimated standard deviations from the control mean

Confidence level =   0.95

          BMD =   4.33789

        BMDL =   2.82726

Linear Model with 0.95 Confidence Level
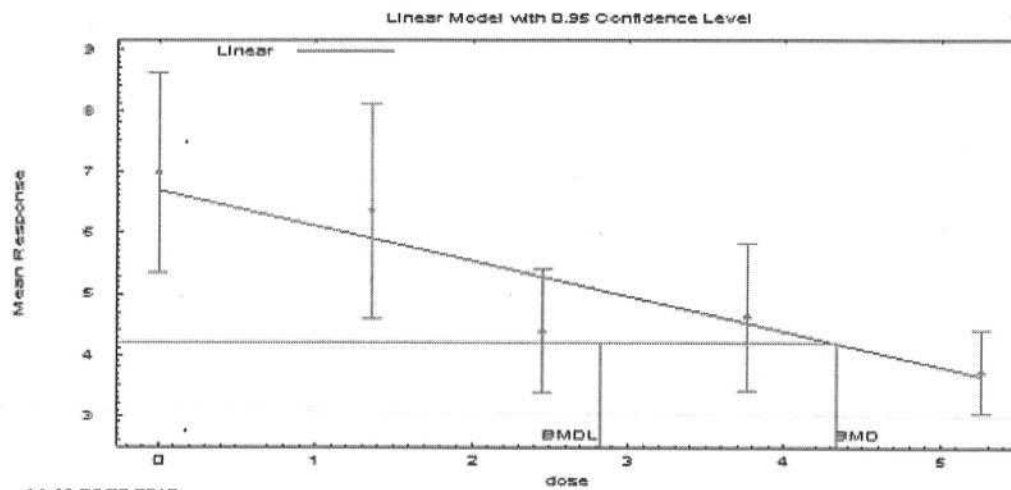
Attachment B

Alternative BMD Models and Modeling Approaches

BMD Modeling Results

As described in the body of the report, effects on reduced WBC and lymphocytes were modeled in the U.S. EPA BMDS after log transformation of dose. The linear, power, and exponential models were all found to provide reasonable fits to the data. In fact, the power model provided the same results as the linear model. Among the submodels that are run simultaneously with the exponential model, submodels 2 and 4 gave acceptable fits to the data (**Table B1**). The output from the BMDS for the continuous models in **Table B1** are included at the end of this discussion. As already described above, the results of the linear model were chosen as the POD because the linear model provided the best fit overall and is also the most parsimonious of the models tested.

Another approach to modeling continuous data is to dichotomize the data and model the incidence of adverse effects (U.S. EPA, 2000). This approach was explored because there is uncertainty as to whether the observed changes in white blood cell counts truly represent an adverse health effect. Typically, reference ranges for hematological values are considered normal when they fall within ± 2 standard deviations from the mean value (Sucklow et al., 2006). As such, historical control mean and standard deviation values were also used to set cutoffs for scoring the incidence of low cell count in individual animals in the HLS sulfolane study. Therefore, two cutoff values were chosen for the WBC and lymphocytes datasets: (-) 2 standard deviations from the mean using the historical HLS data, and (-) 2 standard deviations from the mean in the HLS concurrent controls. Using (-) 2 standard deviations from the concurrent and historical HLS datasets resulted in identical incidences for reduced WBC and lymphocytes, which is not unexpected given that lymphocytes comprise the vast majority of white blood cells. The $BMDL_{10}$ for reduction in white blood cells in female rats exposed to sulfolane in drinking water was determined to be 21.8 mg/kg/day (**Table B1**).

Although the dichotomization approach takes into account the biological relevance of the reduced white blood cell count through scoring individual animals as having abnormally low cell counts, the small sample size makes this approach less robust due to the loss of information and statistical power after converting the data to incidence. Among the

continuous models that fit the WBC data, submodels 2 and 4 gave slightly higher p-values and slightly lower[3] Akaike's Information Criterion (AIC) values than the linear model, but the differences were not enough to clearly establish which model has the better fit (**Table B1**). In contrast, the scaled residuals closest to the BMD were about 2-fold lower in the linear model than the exponential models. Furthermore, in evaluating the EPA BMD modeling results of benzene on lymphocytes using this modeling approach, the EPA stated that the "the linear model was selected because it is the most parsimonious" (U.S. EPA, 2002a). Given that the reduction in leukocytes were not clearly adverse effects and that sulfolane does not appear to be genotoxic, the BMDL values from the parsimonious linear models were selected as the POD as was done in the assessment for benzene (U.S. EPA 2002a).

---

[3] lower AIC values are better.

Table B1. Summary of Benchmark Modeling Results

Model Parameter	BMD Modeling Results			
	Continuous			Dichotomized

WBC

Model	Linear[1]	Exponential		LogLogistic
Submodel		M2	M4	
p-values	0.1677	0.1755	0.1755	0.54
scaled residual	0.168	0.3819	0.3819	0.322
AIC	113.08	112.97	112.97	28.6
BMD^2, ln(dose+1)	4.22	3.81	3.81	NA
BMDL, ln(dose+1)	2.78	2.23	1.78	NA
BMD	67.03	44.15	44.15	68.9
BMDL	15.12	8.30	4.93	21.8

Lymphocytes

Model	Linear[1]	Exponential		LogLogistic
Submodel		M2	M4	
p-values	0.158	0.1678	0.1678	0.54
scaled residual	0.232	0.4715	0.4715	0.322
AIC	102.61	102.46	102.46	28.6
BMD^2, ln(dose+1)	4.34	3.86	3.86	NA
BMDL, ln(dose+1)	2.83	2.19	1.68	NA
BMD	75.71	46.28	46.47	68.9
BMDL	15.95	7.94	4.37	21.8

[1] identical values were also obtained with the BMDS Power Model
[2] Because the doses were log transformed, the BMD and BMDL values reported in the BMD software output were ln(dose + 1) and were manually converted back to arithmetic scale for reporting BMD and BMDL

```
==================================================================
            Exponential Model. (Version: 1.61;  Date: 7/24/2009)
            Input Data File: C:\USEPA\BMDS21\Data\exp_HLS-WBC_logdose-historicalHLS-
SD_Setting.(d)
            Gnuplot Plotting File:
                                          Wed Jun 02 13:47:59 2010
==================================================================
```

BMDS Model Run
~~~~~~

   The form of the response function by Model:
      Model 2:   Y[dose] = a * exp{sign * b * dose}
      Model 3:   Y[dose] = a * exp{sign * (b * dose)^d}
      Model 4:   Y[dose] = a * [c-(c-1) * exp{-b * dose}]
      Model 5:   Y[dose] = a * [c-(c-1) * exp{-(b * dose)^d}]

   Note: Y[dose] is the median response for exposure = dose;
         sign = +1 for increasing trend in data;
         sign = -1 for decreasing trend.

      Model 2 is nested within Models 3 and 4.
      Model 3 is nested within Model 5.
      Model 4 is nested within Model 5.

   Dependent variable = Response
   Independent variable = Dose
   Data are assumed to be distributed: normally
   Variance Model: exp(lnalpha +rho *ln(Y[dose]))
   The variance is to be modeled as Var(i) = exp(lalpha + log(mean(i)) * rho)

   Total number of dose groups = 5
   Total number of records with missing values = 0
   Maximum number of iterations = 250
   Relative Function Convergence has been set to: 1e-008
   Parameter Convergence has been set to: 1e-008

   MLE solution provided: Exact
```

<center>Initial Parameter Values</center>

| Variable | Model 2 | Model 3 | Model 4 | Model 5 |
|---|---|---|---|---|
| lnalpha | -4.46856 | -4.46856 | -4.46856 | -4.46856 |
| rho | 3.12885 | 3.12885 | 3.12885 | 3.12885 |
| a | 4.59626 | 4.59626 | 8.3685 | 8.3685 |
| b | 0.111231 | 0.111231 | 0.140286 | 0.140286 |
| c | -- | -- | 0.108502 | 0.108502 |
| d | -- | 1 | -- | 1 |

<center>Parameter Estimates by Model</center>

| Variable | Model 2 | Model 3 | Model 4 | Model 5 |
|------|------|------|------|------|
| lnalpha | -4.38524 | -4.46381 | -4.38524 | - |
| | | | | 4.27557 |
| rho | 3.07751 | 3.12036 | 3.07751 | |
| | | | | 3.01475 |
| a | 8.10467 | 7.97624 | 8.10467 | |
| | | | | 8.00292 |
| b | 0.110789 | 0.114998 | 0.110789 | |
| | | | | 0.281389 |
| c | -- | -- | 0 | |
| | | | | 0.481718 |
| d | -- | 1.10869 | -- | 1.47486 |

### Table of Stats From Input Data

| Dose | N | Obs Mean | Obs Std Dev |
|------|------|------|------|
| 0 | 10 | 7.97 | 2.626 |
| 1.361 | 10 | 7.63 | 2.653 |
| 2.451 | 9 | 5.41 | 1.392 |
| 3.761 | 9 | 5.53 | 1.756 |
| 5.258 | 10 | 4.54 | 1.109 |

### Estimated Values of Interest

| Model | Dose | Est Mean | Est Std | Scaled Residual |
|------|------|------|------|------|
| 2 | 0 | 8.105 | 2.793 | -0.1525 |
| | 1.361 | 6.97 | 2.215 | 0.942 |
| | 2.451 | 6.177 | 1.839 | -1.252 |
| | 3.761 | 5.343 | 1.471 | 0.3819 |
| | 5.258 | 4.526 | 1.14 | 0.03795 |
| 3 | 0 | 7.976 | 2.739 | -0.007205 |
| | 1.361 | 7.018 | 2.244 | 0.862 |
| | 2.451 | 6.239 | 1.867 | -1.332 |
| | 3.761 | 5.374 | 1.479 | 0.3161 |
| | 5.258 | 4.5 | 1.121 | 0.114 |
| 4 | 0 | 8.105 | 2.793 | -0.1525 |
| | 1.361 | 6.97 | 2.215 | 0.942 |
| | 2.451 | 6.177 | 1.839 | -1.252 |
| | 3.761 | 5.343 | 1.471 | 0.3819 |
| | 5.258 | 4.526 | 1.14 | 0.03795 |
| 5 | 0 | 8.003 | 2.711 | -0.0384 |
| | 1.361 | 7.109 | 2.268 | 0.7268 |
| | 2.451 | 6.182 | 1.837 | -1.261 |
| | 3.761 | 5.254 | 1.437 | 0.577 |
| | 5.258 | 4.553 | 1.159 | -0.03603 |

Other models for which likelihoods are calculated:

Model A1:        $Y_{ij} = Mu(i) + e(ij)$
         $Var\{e(ij)\} = Sigma^2$

Model A2:        $Y_{ij} = Mu(i) + e(ij)$
         $Var\{e(ij)\} = Sigma(i)^2$

Model A3:        $Y_{ij} = Mu(i) + e(ij)$

```
                    Var{e(ij)} = exp(lalpha + log(mean(i)) * rho)

    Model   R:          Yij = Mu + e(i)
                 Var{e(ij)} = Sigma^2
```

<div align="center">Likelihoods of Interest</div>

Model	Log(likelihood)	DF	AIC
A1	-55.27087	6	122.5417
A2	-49.84968	10	119.6994
A3	-50.01255	7	114.0251
R	-65.0528	2	134.1056
2	-52.48727	4	112.9745
3	-52.45111	5	114.9022
4	-52.48727	4	112.9745
5	-52.36868	6	116.7374

Additive constant for all log-likelihoods =   -44.11. This constant added to the
above values gives the log-likelihood including the term that does not
depend on the model parameters.

<div align="center">Explanation of Tests</div>

Test 1: Does response and/or variances differ among Dose levels? (A2 vs. R)
Test 2: Are Variances Homogeneous? (A2 vs. A1)
Test 3: Are variances adequately modeled? (A2 vs. A3)
Test 4: Does Model 2 fit the data? (A3 vs. 2)

Test 5a: Does Model 3 fit the data? (A3 vs 3)
Test 5b: Is Model 3 better than Model 2? (3 vs. 2)

Test 6a: Does Model 4 fit the data? (A3 vs 4)
Test 6b: Is Model 4 better than Model 2? (4 vs. 2)

Test 7a: Does Model 5 fit the data? (A3 vs 5)
Test 7b: Is Model 5 better than Model 3? (5 vs. 3)
Test 7c: Is Model 5 better than Model 4? (5 vs. 4)

<div align="center">Tests of Interest</div>

Test	-2*log(Likelihood Ratio)	D. F.	p-value
Test 1	30.41	8	0.0001791
Test 2	10.84	4	0.02839
Test 3	0.3257	3	0.9551
Test 4	4.949	3	0.1755
Test 5a	4.877	2	0.08729
Test 5b	0.07233	1	0.788
Test 6a	4.949	3	0.1755
Test 6b	0	0	N/A
Test 7a	4.712	1	0.02995
Test 7b	0.1649	1	0.6847
Test 7c	0.2372	2	0.8882

The p-value for Test 1 is less than .05. There appears to be a
difference between response and/or variances among the dose
levels, it seems appropriate to model the data.

The p-value for Test 2 is less than .1. A non-homogeneous
variance model appears to be appropriate.

The p-value for Test 3 is greater than .1. The modeled
variance appears to be appropriate here.

The p-value for Test 4 is greater than .1. Model 2 seems
to adequately describe the data.

The p-value for Test 5a is less than .1. Model 3 may not adequately
describe the data; you may want to consider another model.

The p-value for Test 5b is greater than .05. Model 3 does
not seem to fit the data better than Model 2.

The p-value for Test 6a is greater than .1. Model 4 seems
to adequately describe the data.

Degrees of freedom for Test 6b are less than or equal to 0.
The Chi-Square test for fit is not valid.

The p-value for Test 7a is less than .1. Model 5 may not adequately
describe the data; you may want to consider another model.

The p-value for Test 7b is greater than .05. Model 5 does
not seem to fit the data better than Model 3.

The p-value for Test 7c is greater than .05. Model 5 does
not seem to fit the data better than Model 4.

Benchmark Dose Computations:

  Specified Effect = 1.000000

         Risk Type = Estimated standard deviations from control

  Confidence Level = 0.950000

                BMD and BMDL by Model

        Model   BMD   BMDL
        -------   -------------   -----------
          2   3.81396   2.22988
          3   3.98291   2.24415
          4   3.81396   1.78368
          5   3.69718   1.40591

Lymphocytes, Log Dose, Historical HLS SD - Exponential

```
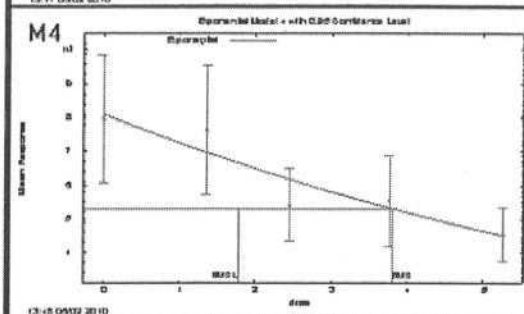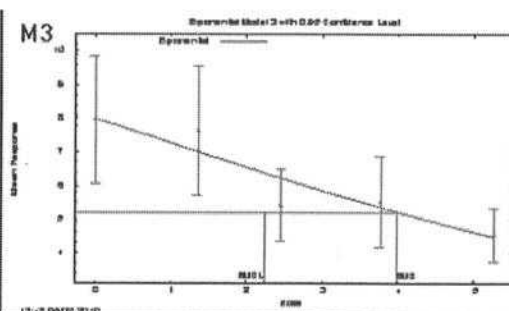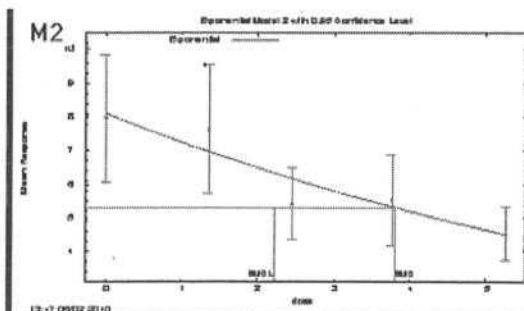==========================================================================
          Exponential Model. (Version: 1.61;  Date: 7/24/2009)
          Input Data File: C:\USEPA\BMDS21\Data\exp_HLS-Lymphocytes-logdose-
historicalHLS-SD_Setting.(d)
          Gnuplot Plotting File:
                                      Wed Jun 02 14:55:55 2010 ·
==========================================================================
```

BMDS Model Run

~~~~~~
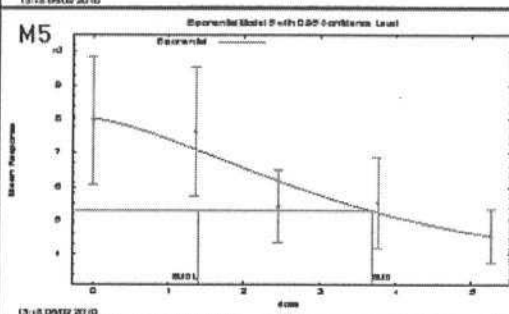
   The form of the response function by Model:
      Model 2:   Y[dose] = a * exp{sign * b * dose}
      Model 3:   Y[dose] = a * exp{sign * (b * dose)^d}
      Model 4:   Y[dose] = a * [c-(c-1) * exp{-b * dose}]
      Model 5:   Y[dose] = a * [c-(c-1) * exp{-(b * dose)^d}]

   Note: Y[dose] is the median response for exposure = dose;
         sign = +1 for increasing trend in data;
         sign = -1 for decreasing trend.

      Model 2 is nested within Models 3 and 4.
      Model 3 is nested within Model 5.
      Model 4 is nested within Model 5.

   Dependent variable = Response
   Independent variable = Dose
   Data are assumed to be distributed: normally
   Variance Model: exp(lnalpha +rho *ln(Y[dose]))
   The variance is to be modeled as Var(i) = exp(lalpha + log(mean(i)) * rho)

   Total number of dose groups = 5
   Total number of records with missing values = 0
   Maximum number of iterations = 250
   Relative Function Convergence has been set to: 1e-008
   Parameter Convergence has been set to: 1e-008

   MLE solution provided: Exact

                     Initial Parameter Values

        Variable   Model 2   Model 3   Model 4   Model 5
        --------   -------   -------   -------   -------
        lnalpha   -3.80574   -3.80574   -3.80574   -
3.80574
        rho   2.92924   2.92924   2.92924
2.92924
        a   3.75106   3.75106   7.329
7.329
        b   0.120754   0.120754   0.208881
0.208881
        c   --   --   0.254469
0.254469
        d   --   1   --
1
```

```
                      Parameter Estimates by Model

        Variable        Model 2         Model 3         Model 4         Model 5
        --------        -------         -------         -------         -------
        lnalpha        -3.90323        -3.99572        -3.90323           -
3.85997           .
         rho            2.98476         3.04094         2.98476
2.95686
          a             6.9219          6.82651         6.9219
6.84835
          b             0.118982        0.121699        0.118982
0.255061
          c               --              --               0
0.408798
          d               --            1.08466           --            1.34213
```

```
                  Table of Stats From Input Data

      Dose       N        Obs Mean      Obs Std Dev
      -----     ---      ----------    --------------
          0      10         6.98           2.29
      1.361      10         6.36           2.452
      2.451       9         4.39           1.308
      3.761       9         4.63           1.564
      5.258      10         3.73           0.941
```

```
                    Estimated Values of Interest

      Model      Dose      Est Mean      Est Std     Scaled Residual
      -----     ------    ----------    ---------    ----------------
        2          0         6.922        2.549          0.07208
                 1.361       5.887        2.002          0.7471
                 2.451       5.171        1.649          -1.42
                 3.761       4.425        1.307          0.4715
                 5.258       3.703        1.002          0.08592
        3          0         6.827        2.516          0.1929
                 1.361       5.921        2.027          0.6844
                 2.451       5.215        1.671          -1.482
                 3.761       4.448        1.312          0.417
                 5.258       3.686        0.9859         0.1398
        4          0         6.922        2.549          0.07208
                 1.361       5.887        2.002          0.7471
                 2.451       5.171        1.649          -1.42
                 3.761       4.425        1.307          0.4715
                 5.258       3.703        1.002          0.08592
        5          0         6.848        2.496          0.1668
                 1.361       5.979        2.042          0.59
                 2.451       5.177        1.65           -1.431
                 3.761       4.372        1.285          0.6022
                 5.258       3.719        1.012          0.03536
```

Other models for which likelihoods are calculated:

    Model A1:   Yij = Mu(i) + e(ij)
              Var{e(ij)} = Sigma^2

    Model A2:   Yij = Mu(i) + e(ij)
              .Var{e(ij)} = Sigma(i)^2

```
Model A3:        Yij = Mu(i) + e(ij)
         Var{e(ij)} = exp(lalpha + log(mean(i)) * rho)

Model  R:        Yij = Mu + e(i)
         Var{e(ij)} = Sigma^2
```

<div align="center">Likelihoods of Interest</div>

Model	Log(likelihood)	DF	AIC
A1	-50.12088	6	112.2418
A2	-44.44769	10	108.8954
A3	-44.70446	7	103.4089
R	-60.31932	2	124.6386
2	-47.2319	4	102.4638
3	-47.21004	5	104.4201
4	-47.2319	4	102.4638
5	-47.16859	6	106.3372

Additive constant for all log-likelihoods =   -44.11. This constant added to the
above values gives the log-likelihood including the term that does not
depend on the model parameters.

<div align="center">Explanation of Tests</div>

Test 1: Does response and/or variances differ among Dose levels? (A2 vs. R)
Test 2: Are Variances Homogeneous? (A2 vs. A1)
Test 3: Are variances adequately modeled? (A2 vs. A3)
Test 4: Does Model 2 fit the data? (A3 vs. 2)

Test 5a: Does Model 3 fit the data? (A3 vs 3)
Test 5b: Is Model 3 better than Model 2? (3 vs. 2)

Test 6a: Does Model 4 fit the data? (A3 vs 4)
Test 6b: Is Model 4 better than Model 2? (4 vs. 2)

Test 7a: Does Model 5 fit the data? (A3 vs 5)
Test 7b: Is Model 5 better than Model 3? (5 vs. 3)
Test 7c: Is Model 5 better than Model 4? (5 vs. 4)

<div align="center">Tests of Interest</div>

Test	-2*log(Likelihood Ratio)	D. F.	p-value
Test 1	31.74	8	0.0001035
Test 2	11.35	4	0.02294
Test 3	0.5135	3	0.9159
Test 4	5.055	3	0.1678
Test 5a	5.011	2	0.08163
Test 5b	0.04371	1	0.8344
Test 6a	5.055	3	0.1678
Test 6b	1.421e-014	0	N/A
Test 7a	4.928	1	0.02642
Test 7b	0.0829	1	0.7734
Test 7c	0.1266	2	0.9387

The p-value for Test 1 is less than .05. There appears to be a

difference between response and/or variances among the dose
levels, it seems appropriate to model the data.

The p-value for Test 2 is less than .1. A non-homogeneous
variance model appears to be appropriate.

The p-value for Test 3 is greater than .1. The modeled
variance appears to be appropriate here.

The p-value for Test 4 is greater than .1. Model 2 seems
to adequately describe the data.

The p-value for Test 5a is less than .1. Model 3 may not adequately
describe the data; you may want to consider another model.

The p-value for Test 5b is greater than .05. Model 3 does
not seem to fit the data better than Model 2.

The p-value for Test 6a is greater than .1. Model 4 seems
to adequately describe the data.

Degrees of freedom for Test 6b are less than or equal to 0.
The Chi-Square test for fit is not valid.

The p-value for Test 7a is less than .1. Model 5 may not adequately
describe the data; you may want to consider another model.

The p-value for Test 7b is greater than .05. Model 5 does
not seem to fit the data better than Model 3.

The p-value for Test 7c is greater than .05. Model 5 does
not seem to fit the data better than Model 4.
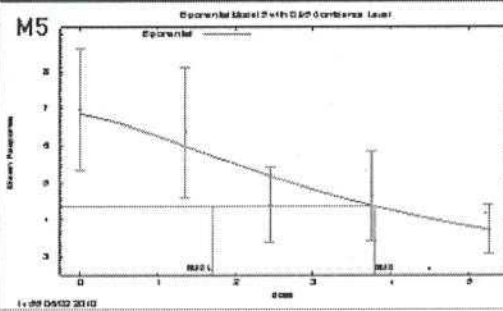

Benchmark Dose Computations:
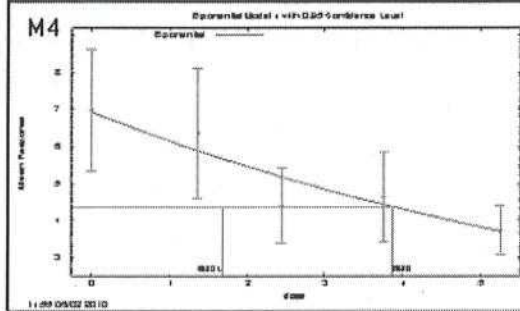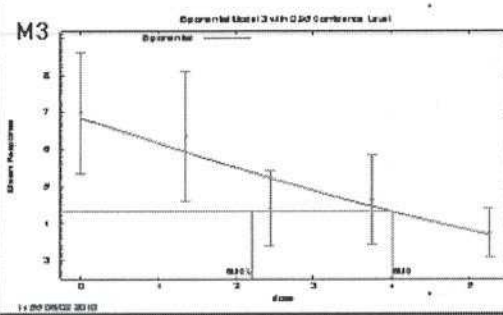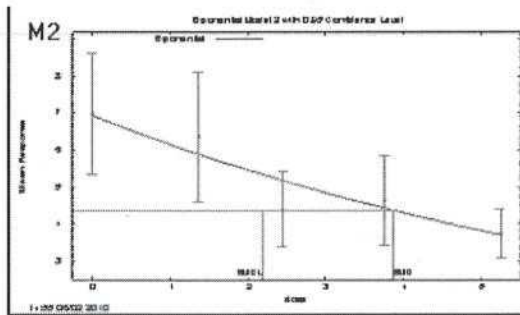
 Specified Effect = 1.000000

        Risk Type = Estimated standard deviations from control

 Confidence Level = 0.950000

                 BMD and BMDL by Model

     Model   BMD   BMDL
     -------   ------------   ----------
        2   3.85985   2.19274
        3   4.01408   2.20163
        4   3.85985   1.68317
        5   3.79761   1.70124

From:	foia@erulemaking.net
Sent:	Tuesday, November 17, 2015 4:30 PM
To:	Bailey, Marcia
Subject:	FOIA Assignment for Request Detail Task for Request EPA-R10-2016-001183

You have been assigned to the FOIA task - Request Detail Task. Additional details for this item are as follows:

- Assigned By: Sharon Buza
- Tracking Number: EPA-R10-2016-001183
- Due Date: 11/18/2015
- Sent Date: 11/17/2015
- Requester: Meredith Weinberg
- Request Track: Simple
- Short Description: research and upload responsive records. Keep track and record time
- Long Description: N/A